A VERIFICATION TASK WITH LATERALIZED TONES AND ACCELERATED SPEECH

Michael A. Nees & Kathryn Best

Lafayette College
Department of Psychology
Oechsle Hall, Easton, PA, USA
[neesm, bestk]@lafayette.edu

ABSTRACT

Research has suggested that the left hemisphere of the brain may be specialized for processing auditory speech, whereas the right hemisphere may be specialized for processing nonspeech auditory stimuli. Due to contralaterality in auditory pathways, this functional specialization has been reflected in behavioral advantages for speech stimuli presented to the right ear and for nonspeech stimuli presented to the left ear. We used a verification task with lateralized presentations of brief tonal stimuli (sonifications) and accelerated speech stimuli (spearcons) to examine performance as a function of the presentation ear and the type of auditory display. The general pattern of results showed that reaction time and accuracy were facilitated when two accelerated speech stimuli were compared to each other. Based on the results of this study, reported effects of left and right ear advantages do not seem to be robust enough to warrant general ergonomic recommendations (i.e., left ear presentation of nonspeech sounds and right ear presentation of speech sounds) for auditory display design.

1. INTRODUCTION

Research has suggested that the left and right hemispheres of the brain may be specialized for processing different properties of auditory stimuli. [1][2]. Zatorre and colleagues reviewed evidence to suggest that the left hemisphere is specialized for resolving fine temporal differences in sounds, whereas the right hemisphere is specialized for fine-grained analysis of spectral acoustic properties. As such, the left hemisphere may be specialized for processing auditory speech stimuli, whereas the right hemisphere may be specialized for processing musical, tonal, and nonspeech stimuli (also see [3]).

The auditory system has both ipsilateral and contralateral cochlear connections that converge early in the auditory pathway and most notably serve sound localization functions (see, e.g., [4]). In other words, hemispheric input from the periphery is not exclusive to a particular ear, although research has suggested that contralateral input may be dominant [5][6]. Several studies have demonstrated a *right ear advantage* (REA) for processing verbal information, and a parallel *left ear advantage* (LEA) has been shown for processing nonspeech auditory and tonal stimuli (see [6]–[9]). The possible impact of hemispheric functional specialization on ergonomics decisions about peripheral presentation of sounds has not been fully explored.

Research to date has not considered the possible applications of the REA and LEA to research on sonification and auditory displays. In particular, researchers have not examined the possibility that sonifications and other nonspeech auditory displays might best be processed with lateralized, left ear presentations. Similarly, performance with speech auditory displays could perhaps be optimized with lateralized, right ear presentations. If confirmed across a robust variety of auditory display scenarios, the REA and LEA could be translated to simple heuristics for improving performance with auditory displays.

The current study examined performance with lateralized verbal and tonal stimuli using a speeded verification task. Verification tasks have been used to examine cognitive processes, including encoding strategies in working memory (see [10]). In the general form of a verification task, participants are presented with an initial stimulus to study or remember. The initial stimulus usually conveys the state of a simple, binary scenario. In the current study and other recent work [11][12], we have used stimuli that depicted a fictional stock price state as simply increasing or decreasing, but the particular application domain was of less importance to the current study than establishing a simple tonal stimulus set for the verification. After presentation of the initial stimulus, a second stimulus is presented for comparison. The second stimulus depicts one of the same possible binary states (e.g., increasing or decreasing). The participants' task is to judge as quickly as possible whether the states depicted in the first (study) and second (verification) stimuli match or mismatch.

In the current study, we used a verification task with lateralized (fully left ear or fully right ear) presentations of brief tonal and accelerated speech stimuli to depict the increasing or decreasing state of a fictional stock price. The tonal and accelerated speech stimuli were equivalent to brief sonifications and spearcons, respectively (for a review and description of auditory display design elements, see [13]). The study was designed to examine potential LEAs for nonspeech audio sonifications and REAs for (presumably) verbal accelerated speech auditory stimuli. As such, we predicted that verification response times would be fastest when: a) tonal verification stimuli were presented to the left ear; and b) accelerated speech verification stimuli were presented to the right ear. Previous research [11][12] has also suggested that verification times can sometimes be facilitated when the format of the verification stimulus matches the format of the study stimulus (i.e. tones are compared to tones; speech is compared to speech), so planned analyses also examined this possibility.

2. METHOD

2.1. Participants

Participants (N = 31, M age = 19.81, SD = 0.91 years, 23 females) were recruited from undergraduate psychology courses at Lafayette College and were compensated with course credit. At the beginning of the study, participants self-reported handedness and completed a demographics questionnaire that include three musical experience questions that queried participants about the number of years they had played a musical instrument, their number of years of formal musical training, and their number of years of experience with reading musical notation.

2.2. Apparatus and Stimuli

Presentations of stimuli and data collection were accomplished with a program written in Adobe Director. Audio stimuli were presented with Sony MDR-V6 headphones.

2.2.1. Tonal stimuli

Tonal stimuli were lateralized versions of the tonal stimuli described in [11]. The stimuli consisted of two discrete, 100 ms duration notes with 10 ms onset and offset ramps. The notes C4 (262 Hz) and C5 (523 Hz) were used, and the notes were synthesized with the MIDI piano instrument. An increase in the stimulus state was represented with C4 followed by C5, and a decrease was represented with C5 followed by C4. All tonal stimuli were 200 ms in duration. Lateralized versions of the increasing and decreasing stimuli were created with Audacity software to locate the sounds entirely in the left or right headphone channel, so a total of four tonal stimuli were used.

2.2.2. Accelerated Speech Stimuli

Accelerated speech stimuli were lateralized versions of the tonal stimuli described in [11]. Acceleration of speech was required for auditory stimuli (tones and speech) to be equal in duration, and the resulting stimuli were effectively spearcons [14]. The stimuli consisted of the spoken words "increase" and "decrease." The words were created as WAV files with the online AT&T Labs TTS demonstration and used a female voice (*Crystal, US English*). The TTS WAV files were compressed to 200 ms duration with the Audacity software's *change tempo without changing pitch* function. Lateralized versions of the increasing and decreasing stimuli were created with Audacity software to locate the sounds entirely in the left or right headphone channel, so a total of four speech stimuli were used.

2.3. Procedure

At the beginning of the procedure, the experimenter verified that the left and right headphone channels were positioned correctly. A trial (see Figure 1) consisted of the presentation of a study stimulus (one of the eight possible stimuli), a 3000 ms delay, and the presentation of a verification stimulus (again, one of the eight possible stimuli). Participants' task was to compare the state depicted in the study stimulus (increasing or decreasing) to the state depicted in the verification stimulus and to respond as quickly as possible to indicate whether the states did or did not match. Both study response times and verification response times were recorded. Responses were matched to the "Z" and "?" keys on the computer keyboard, and the mapping of match versus mismatch responses to keys was counterbalanced across participants. Participants received feedback about their verification reaction time following every trial. Participants experienced a total of 128 trials in two blocks. Every possible pairing of the eight possible study and verification stimuli was presented twice across the entire study.

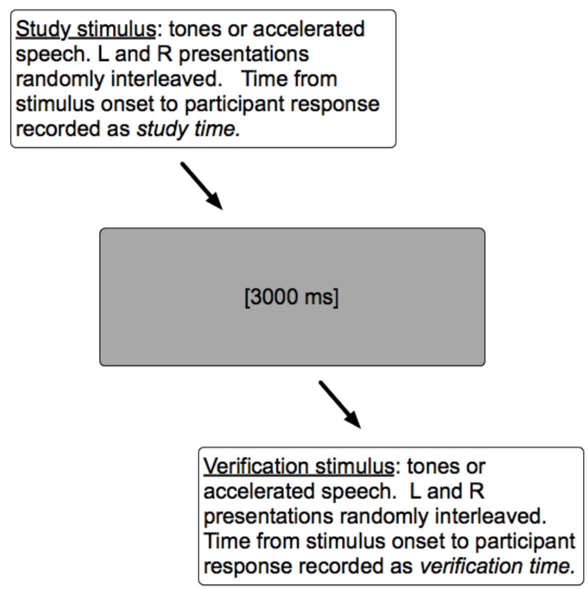

Figure 1: Structure of an experimental trial.

3. RESULTS

Data from one participant was removed from all analyses after the participant self-reported that he was ambidextrous, because the REA effect may potentially be limited to right handed people [15]. For all response time measures, outliers were defined as scores that fell outside of three times the interquartile range for a given data point, and these values were removed from all analyses. All response time measures are reported in milliseconds. Both verification times and the proportion of correct responses were analyzed with 2 (study stimulus type) X 2 (verification stimulus types) X 2 (verification ear) repeated measures analyses of variance (ANOVAs).

For verification times, there were no main effects of either study stimulus type, $F(1, 29) = 2.49$, $p = .13$, or verification ear, $F(1, 29) = 0.32$, $p = .58$. The main effect of verification stimulus type was significant, $F(1, 29) = 39.33$, $p < .001$, η_p^2

= .58; participants responded more slowly when the verification stimulus was a tone (M = 1183.91, SE = 55.49) as compared to accelerated speech (M = 1000.92, SE = 42.12). There was a significant interaction of study stimulus type with verification stimulus type, $F(1, 29) = 24.22$, $p < .001$, $\eta_p^2 = .46$. Simple effects at each level of study stimulus type showed that, when the study stimulus was speech, participants were faster to respond to a speech verification stimulus (M = 940.70, SE = 38.4) as compared to tones (M = 1273.71, SE = 70.68), $p < .001$. When the study stimulus was tones, there was not a significant difference between speech (M = 1061.14, SE = 48.95) and tonal (M = 1094.11, SE = 46.05) verification stimuli, p = .36. No other omnibus interactions were statistically significant (ps = .20 - .85)(see Figure 2).

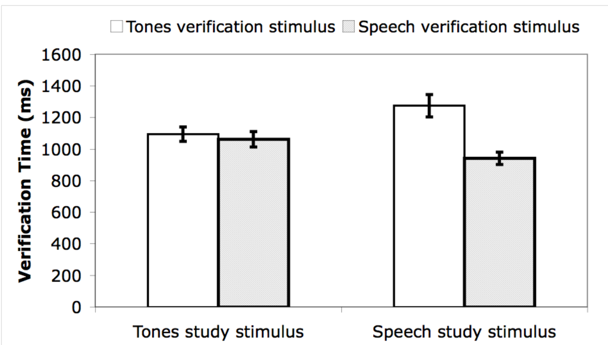

Figure 2: Mean verification time as a function of study stimulus format and verification stimulus format (collapsed across verification ear). Error bars represent standard error of the mean.

For the proportion of correct responses, there was no main effect of study stimulus type, $F(1, 29) = 2.53$, p = .12. The main effect of verification stimulus type was significant, $F(1, 29) = 6.48$, p = .016, $\eta_p^2 = .18$; participants made significantly fewer correct responses when the verification stimulus was a tone (M = .823, SE = .026) as compared to accelerated speech (M = .849, SE = .020). The main effect of verification ear also was significant, $F(1, 29) = 5.22$, p = .03, $\eta_p^2 = .15$; participants made significantly fewer correct responses when the verification stimulus was presented to the left ear (M = .83, SE = .023) as compared to the right ear (M = .847, SE = .022). There was a significant interaction of study stimulus type with verification stimulus type, $F(1, 29) = 10.18$, p = .003, $\eta_p^2 = .26$. Simple effects at each level of study stimulus type showed that, when the study stimulus was speech, participants made a higher percentage of correct responses to speech verification stimuli (M = .891, SE = .022) as compared to tones (M = .807, SE = .025), p = .002. When the study stimulus was tones, there was not a significant difference between correct responses for speech (M = .819, SE = .026) and tonal (M = .838, SE = .028) verification stimuli, p = .24 (see Figure 3). No other omnibus interactions were statistically significant (ps = .56 - .98).

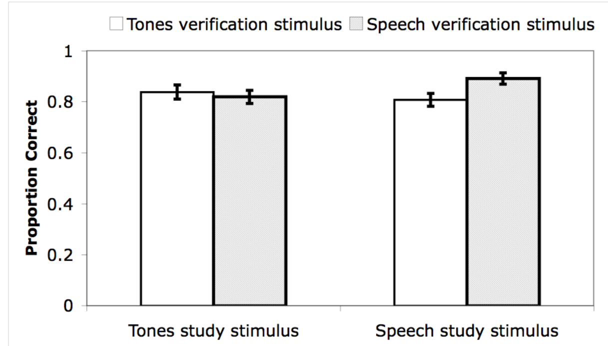

Figure 3: Mean proportion correct as a function of study stimulus format and verification stimulus format (collapsed across verification ear). Error bars represent standard error of the mean.

An exploratory analysis examined study times—the amount of time from the study stimulus onset until the participant indicated they were ready to continue. The 2 (study stimulus type) X 2 (study ear) ANOVA showed only a significant main effect of study stimulus type, $F(1, 29) = 57.75$, $p < .001$, η_p^2 = .67. Participants spent less time studying speech (M = 1051.50, SE = 43.84) as compared to tones (M = 1402.74, SE = 74.63). The other omnibus effects were not significant, ps = .31 and .47.

A final exploratory analysis examined the correlations of the sum of the three music experience questions with the verification time grand mean, $r(28) = -.35$, p = .03 and the proportion correct grand mean $r(28) = .48$, p = .004. This pattern suggested that participants with more musical experience tended to perform the verification task faster and more accurately. Though suggestive, these analyses were exploratory and ancillary to the primary objectives of this study, and a larger sample would be warranted to draw strong conclusions about the relationship between musical experience and the tasks used here.

4. DISCUSSION

We did not find evidence to support the hypothesis that the tonal and accelerated speech stimuli would result in LEAs and REAs, respectively. In the literature, some qualifications have emerged to the general pattern of results found in REA and LEA studies. One study showed REAs and LEAs only after several blocks of stimuli were repeated [16]. Another study [17], however, showed that the LEA effect was confined only to early trials. Research [18] showed that LEA effects for tonal stimuli were attenuated as time elapsed and disappeared altogether around 5 s following stimulus presentation. Another study suggested that, whereas three-tone stimuli with no frequency changes showed an LEA, a single frequency change in the sequence resulted in practically no difference between ears, and two frequency changes shifted the advantage to the right ear [19]. The stimuli of the current study used a single frequency change. The 200 ms duration of our stimuli appear to be brief as compared to stimulus durations that have demonstrated REA and LEA effects.

The general pattern of results showed that verification was facilitated when a speech study stimulus was paired with a speech verification stimulus. Previous research demonstrated the same pattern of facilitation with monaural presentation (of the same sound to both ears) for both the accelerated speech stimuli [11] and tonal stimuli [11][12] used here. With lateralized stimuli, the effect was replicated for accelerated speech, but not for tones.

Two additional findings warrant further research. First, a small but significant effect showed that participants made significantly fewer correct responses when the verification stimulus was presented to the left ear. Further research is required to replicate this effect and expand upon its potential implications. Second, exploratory correlations suggested that participants with more musical experience performed better on the verification task. The relationships between musical experience, musical ability, and performance with auditory displays have not been clearly established (for a discussion, see [13]), and more research in this area could provide valuable insights on the selection and training of auditory display users.

Based on the results of this study, the LEA and REA phenomena do not seem to be robust enough to warrant general ergonomic recommendations (i.e., left ear presentation of nonspeech sounds and right ear presentation of speech sounds) for auditory display design. REAs and LEAs are influenced by a variety of factors, including stimulus properties, task dependencies, and cognitive factors [20]. Further research is needed to determine if REAs and LEAs could be demonstrated with other types of auditory displays and leveraged to the benefit of human listeners. For example, the task used here required perceptual processing of accelerated speech and tones; tasks that require deeper semantic processing of the meaning of stimuli should be explored in future research. Our results showed that comparisons of accelerated speech stimuli to other accelerated speech stimuli are more easily accomplished than comparisons across tones and speech formats.

5. REFERENCES

[1] R. J. Zatorre, P. Belin, and V. B. Penhune, "Structure and function of auditory cortex: Music and speech," *Trends in Cognitive Sciences*, vol. 6, no. 1, pp. 37–46, 2002.

[2] D. L. Molfese, R. B. Freeman Jr, and D. S. Palermo, "The ontogeny of brain lateralization for speech and nonspeech stimuli," *Brain and Language*, vol. 2, pp. 356–368, 1975.

[3] R. Cohn, "Differential cerebral processing of noise and verbal stimuli," *Science*, vol. 172, no. 3983, pp. 599–601, 1971.

[4] R. Y. Litovsky, B. J. Fligor, and M. J. Tramo, "Functional role of the human inferior colliculus in binaural hearing," *Hearing Research*, vol. 165, no. 1, pp. 177–188, 2002.

[5] C. Pantev, B. Ross, P. Berg, T. Elbert, and B. Rockstroh, "Study of the human auditory cortices using a whole-head magnetometer: Left vs. right hemisphere and ipsilateral vs. contralateral stimulation," *Audiology and Neuro-Otology*, vol. 3, no. 2–3, pp. 183–190, 1998.

[6] D. Kimura, "Functional asymmetry of the brain in dichotic listening," *Cortex*, vol. 3, no. 2, pp. 163–178, 1967.

[7] H. J. Kallman, "Can expectancy explain reaction time ear asymmetries?," *Neuropsychologia*, vol. 16, no. 2, pp. 225–228, 1978.

[8] M. Spajdel, K. Jariabková, and I. Riecanskỳ, "The influence of musical experience on lateralisation of auditory processing," *Laterality*, vol. 12, no. 6, pp. 487–499, 2007.

[9] L. Hoch and B. Tillmann, "Laterality effects for musical structure processing: A dichotic listening study," *Neuropsychology*, vol. 24, no. 5, p. 661, 2010.

[10] N. N. Mathews, E. B. Hunt, and C. M. MacLeod, "Strategy choice and strategy training in sentence-picture verification," *Journal of Verbal Learning and Verbal Behavior*, vol. 19, no. 5, pp. 531–548, 1980.

[11] M. A. Nees and K. Best, "Modality and encoding strategy effects on a verification task with accelerated speech, visual text, and tones," in *Proceedings of the 19th International Conference on Auditory Display (ICAD2013)*, Lodz, Poland, 2013, pp. 267–274.

[12] M. A. Nees and B. N. Walker, "Flexibility of working memory encoding in a sentence-picture-sound verification task," *Journal of Cognitive Psychology*, vol. 25, no. 7, pp. 800–807, 2013.

[13] B. N. Walker and M. A. Nees, "Theory of sonification," in *Principles of Sonification: An Introduction to Auditory Display*, T. Hermann, A. Hunt, and J. Neuhoff, Eds. Berlin, Germany: Logos Publishing House, 2011, pp. 9–39.

[14] B. N. Walker, J. Lindsay, A. Nance, Y. Nakano, D. K. Palladino, T. Dingler, and M. Jeon, "Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus," *Human Factors*, vol. 55, no. 1, pp. 157–182, 2013.

[15] M. P. Bryden, "Dichotic listening performance, cognitive ability, and cerebral organization.," *Canadian Journal of Psychology*, vol. 40, no. 4, pp. 445–56, 1987.

[16] J. J. Sidtis and M. P. Bryden, "Asymmetrical perception of language and music: Evidence for independent processing strategies," *Neuropsychologia*, vol. 16, no. 5, pp. 627–632, 1978.

[17] H. J. Kallman and M. C. Corballis, "Ear asymmetry in reaction time to musical sounds," *Perception & Psychophysics*, vol. 17, no. 4, pp. 368–370, 1975.

[18] O. Spreen, F. J. Spellacy, and J. R. Reid, "The effect of interstimulus interval and intensity on ear asymmetry for nonverbal stimuli in dichotic listening," *Neuropsychologia*, vol. 8, no. 2, pp. 245–250, Apr. 1970.

[19] Y. Halperin, I. Nachshon, and A. Carmon, "Shift of ear superiority in dichotic listening to temporally patterned nonverbal stimuli," *The Journal of the Acoustical Society of America*, vol. 53, p. 46, 1973.

[20] K. Hugdahl, "Fifty years of dichotic listening research– Still going and going and…," *Brain and Cognition*, vol. 76, no. 2, pp. 211–213, 2011.